Using Entanglement to Increase the Angular Resolution of Very Long Baseline Interferometry

Matthew Thill,* Sam Dolinar,* and Dariush Divsalar*

ABSTRACT. — A key motivation in developing quantum technology has been its identified potential for enhancing sensing capability. We present a methodology for using quantum entanglement to improve several features of Very Long Baseline Interferometry. In particular, we show that sharing a simple bipartite state between two antennas equipped with photon-number-resolving detectors can enable us to beat the shot noise limit in measuring the location of a weak point source, achieving Heisenberg-limited sensing.

I. Introduction: Very Long Baseline Interferometry Setup

Very Long Baseline Interferometry (VLBI) is a technique that uses distantly spaced antennas to measure the brightness distribution of a radio or light source. The concept [1] is based on the source wavefront arriving at one of the antennas with a geometrically induced time delay of τ_g (see Figure 1) so that the two antennas receive signals $V_1(t) = v_1 \cos 2\pi\nu(t - \tau_g)$ and $V_2(t) = v_2 \cos 2\pi\nu t$ for a monochromatic point source signal at frequency ν. The signals $V_1(t)$ and $V_2(t)$ are passed through a correlator, which multiplies and time-averages the signals. In the case of a monochromatic point source, the output correlation is $r(\tau_g) = v_1 v_2 \cos 2\pi\nu\tau_g$, a function of the delay τ_g.

The geometric delay can be expressed in terms of the baseline vector $\mathbf{b}$ between the antennas and the unit direction vector $\mathbf{s}$ to the point source: $\tau_g = \mathbf{b} \cdot \mathbf{s}/c$, where c is the speed of light. Thus, measuring an interference fringe pattern associated with varying τ_g allows us to estimate a source position $\mathbf{s}$. A general monochromatic source

*Communications Architectures and Research Section.

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.
© 2023 California Institute of Technology. Government sponsorship acknowledged.

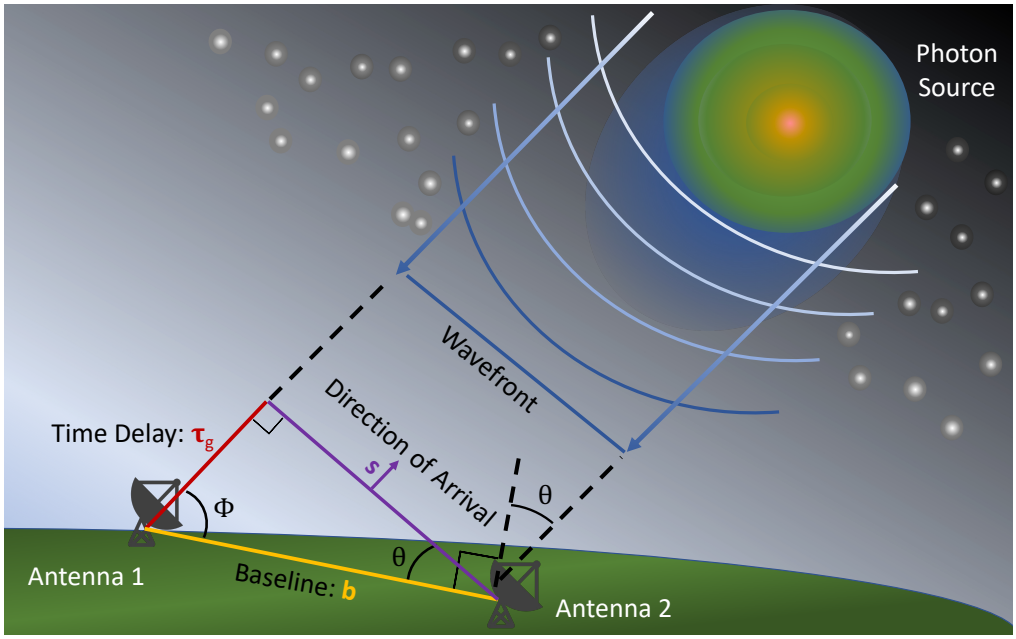


Figure 1. The geometry of a VLBI setup, in which two antennas spaced far apart receive photons from a distant source.

with frequency ν will be defined by a brightness distribution $I(\mathbf{s})$, and our correlator output must be computed by integrating over all position vectors $\mathbf{s}$ or, equivalently, integrating over all solid angle elements $d\Omega$ across the sky. This amounts to measuring the real part of the source *visibility*,

$$V(\mathbf{b}) := \int_S A(\mathbf{s}) I(\mathbf{s}) e^{-2\pi i \nu \mathbf{b}\cdot\mathbf{s}/c} d\Omega \tag{1}$$

where $A(\mathbf{s})$ is the antenna response in direction $\mathbf{s}$. The source brightness distribution $I(\mathbf{s})$ can be recovered as the inverse Fourier transform of $V(\mathbf{b})$. In practice, this necessitates measuring the visibility at a range of different baselines and taking a discrete Fourier transform approximation.

Denoting by Φ the angle between the position vector $\mathbf{s}$ and the baseline vector $\mathbf{b}$, and rewriting $\tau_g = b\cos\Phi/c$ where b is the baseline length, we see that the variational relationship between Φ and τ_g is given by

$$d\Phi = -c \cdot d\tau_g / b \sin\Phi. \tag{2}$$

Thus, the angular resolution of a VLBI system is inversely proportional to the baseline length b [2]. Developing antenna arrays with large baselines is generally a costly endeavor, so alternative methods for improving the angular resolution are of great interest.

II. Sharing Entangled Single-Photon States

Several years ago, Gottesman et al. [3] designed a way to perform long baseline interferometry for very weak sources, where the incoming wavefront could be modeled as a single-photon quantum state being measured by two telescopes. In a two-telescope system, at a given time, a single photon is equally likely to be detected at either the left (L) or the right (R) telescope, and we model its quantum state in the form

$$|\psi\rangle^{\text{src}} = \frac{1}{\sqrt{2}}(|0\rangle_L|1\rangle_R + e^{i\phi}|1\rangle_L|0\rangle_R), \tag{3}$$

where $|0\rangle$ and $|1\rangle$ denote 0- and 1-photon states, respectively, and

$$\phi := 2\pi\nu\tau_g = 2\pi\nu\mathbf{b}\cdot\mathbf{s}/c = 2\pi b\sin\theta/\lambda \tag{4}$$

is the phase offset that the photon accrues when arriving at the left telescope. Here, θ is the direction of arrival of the photon, and λ is its wavelength. The correlation step of the interferometry now involves bringing the light from the two telescopes together at a beamsplitter, which Gottesman et al. [3] argued was difficult for single-photon states over long baselines due to inevitable photon losses and undesired phase shifts arising from imperfections in the optical components, leading to reduced interference fringes.

Gottesman et al. [3] posed the following solution to this problem: Distribute the two modes of a prepared maximally entangled photon state (EPS) to the two telescopes (see Figure 2). This state can then be used to either teleport the received left and right signal states back to the beamsplitter or to perform a special measurement directly at the telescopes. The latter approach is simple to implement: Our distributed EPS takes the form

$$|\psi\rangle^{\text{eps}} = \frac{1}{\sqrt{2}}\left(|0\rangle_L|1\rangle_R + e^{i\delta}|1\rangle_L|0\rangle_R\right), \tag{5}$$

where we use an adjustable delay line to apply a varying phase δ to the photon if it arrives at the left antenna. We combine each telescope's share of the EPS with its received signal state in a 50:50 beamsplitter, and post-select cases in which one photon arrives at each telescope (so that one antenna receives the $|\psi\rangle^{\text{src}}$ photon and the other receives the $|\psi\rangle^{\text{eps}}$ photon). The two telescopes' beamsplitters will see a photon in the same output port with probability $[1 + \text{Re}(V(\mathbf{b})e^{-i\delta})]/2$ and opposite output ports with probability $[1 - \text{Re}(V(\mathbf{b})e^{-i\delta})]/2$, so by varying the delay phase δ, we can measure both the phase and amplitude of the visibility $V(\mathbf{b})$.

III. Increasing the Baseline Using Entangled Multiple-Photon States

We begin by considering a point source so that all astronomical photons have a fixed position vector $\mathbf{s}$ and the phase offset ϕ is constant. Consider the following generalization of the protocol in Gottesman et al. [3]: Instead of distributing single

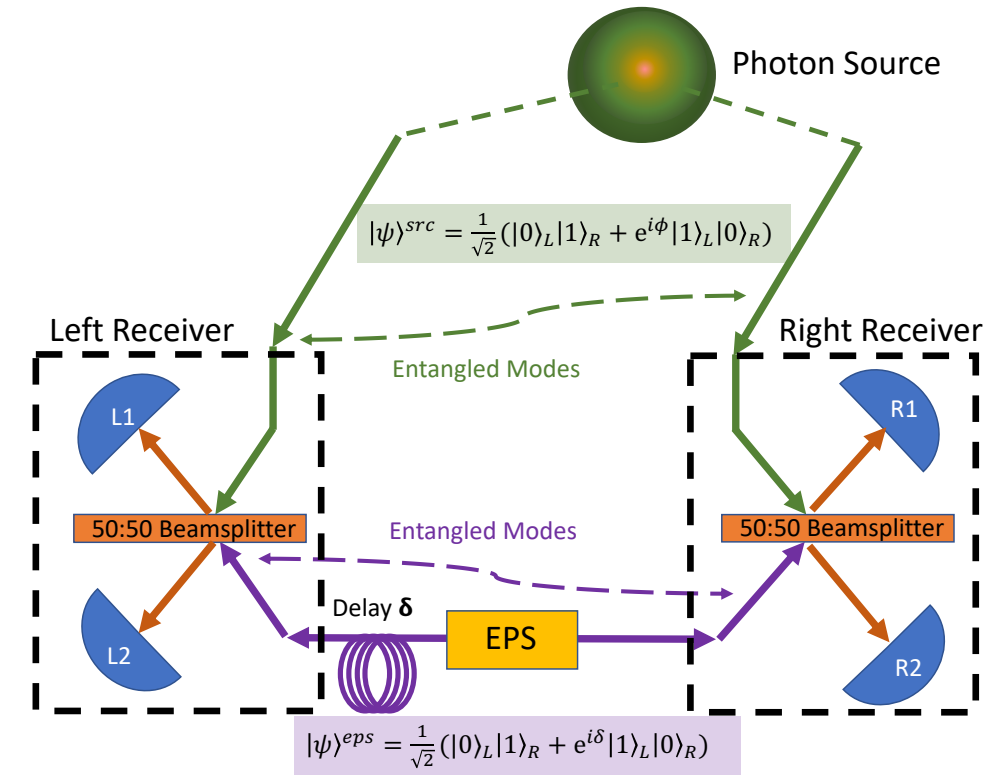


Figure 2. The VLBI receiver discussed in Gottesman et al. [3]. A weak source emits a single photon which can arrive at either the left (L) or the right (R) antenna. Simultaneously, an EPS is shared between the two antennas. The source and EPS modes are joined in a 50:50 beamsplitter at each receiver, with a photon-counting detector at each output. If each receiver receives a single photon click, then the source and EPS photons are indistinguishable, and the configuration of detection events allows for the visibility to be measured by varying an adjustable delay line.

entangled photon pairs, we distribute entangled N-photon states in the form

$$|\Psi_N\rangle^{\text{eps}} = \frac{1}{\sqrt{2}}(|00...0\rangle_L|11...1\rangle_R + e^{iN\delta}|11...1\rangle_L|00...0\rangle_R) \tag{6}$$

$$= \frac{1}{\sqrt{2}}(|N\rangle_R^{\text{eps}} + e^{iN\delta}|N\rangle_L^{\text{eps}}), \tag{7}$$

where we have used the "$|N\rangle$" notation as a shorthand. These are essentially NOON states, and for ease of notation, we will treat them as such. More generally, we may consider the N photons as belonging to distinguishable modes (e.g., temporal), and our following calculations can be adapted accordingly without too much difficulty. Note that in principal, because only one logical entangled qubit (ebit) is encoded in $|\Psi_N\rangle^{\text{eps}}$, it should be possible for the telescope stations to produce this state using local resources and a single shared ebit, perhaps in the form of a single entangled photon pair.

Now, if the two telescopes post-select cases in which each receives N photons (either N photons in a single mode per antenna, or a single photon in each of N different modes per antenna), then the joint state of the EPS and source photons is projected

onto a superposition of the states in which one telescope received the N EPS photons and the other received N source photons:

$$|\Theta_N\rangle_{LR} = \frac{1}{\sqrt{2}}(e^{iN\phi}|N\rangle_L^{\text{src}}|N\rangle_R^{\text{eps}} + e^{iN\delta}|N\rangle_L^{\text{eps}}|N\rangle_R^{\text{src}}). \tag{8}$$

For simplicity's sake, we will first assume that all N EPS photons belong to the same mode, as do the N source photons, as depicted in Figure 3. If we let $\hat{a}_L$ and $\hat{b}_L$ denote the annihilation operators of the source and EPS modes, respectively, at the left telescope, and $\hat{a}_R$ and $\hat{b}_R$ denote the corresponding annihilation operators for the right telescope, then we can reexpress this state as

$$|\Theta_N\rangle_{LR} = \frac{1}{N!\sqrt{2}}(e^{iN\phi}(\hat{a}_L^\dagger)^N(\hat{b}_R^\dagger)^N + e^{iN\delta}(\hat{b}_L^\dagger)^N(\hat{a}_R^\dagger)^N)|0\rangle, \tag{9}$$

where $|0\rangle$ denotes the vacuum state. At each telescope, we input each of the N received source modes and the corresponding EPS mode into the two input arms of a 50:50 beamsplitter. We will denote by $\hat{c}_L$ and $\hat{d}_L$ the two beamsplitter output mode annihilation operators at the left telescope and similarly $\hat{c}_R$ and $\hat{d}_R$ for the right telescope. At the left telescope, we then have the input-output relationships

$$\hat{a}_L^\dagger \mapsto \frac{1}{\sqrt{2}}(\hat{c}_L^\dagger + \hat{d}_L^\dagger), \tag{10}$$

$$\hat{b}_L^\dagger \mapsto \frac{1}{\sqrt{2}}(\hat{c}_L^\dagger - \hat{d}_L^\dagger), \tag{11}$$

and similar relations apply for the right telescope's beamsplitter. Note that we have implicitly normalized the output annihilation operators so that the input-output relationships have real coefficients.

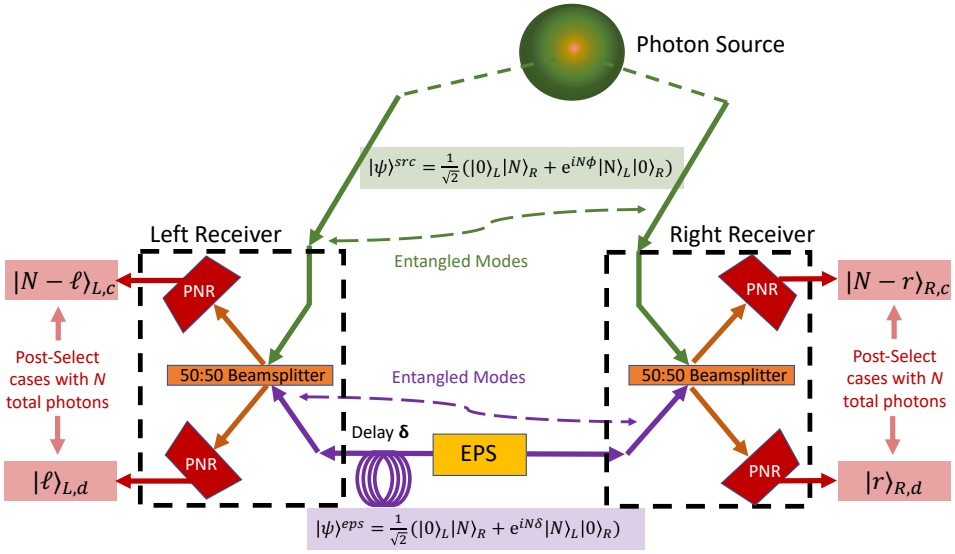


Figure 3. An N-photon extension of the telescope from Gottesman et al. [3]. The EPS is prepared as a superposition of all N photons being sent to either the left or right antenna. We assume photon-number resolution (PNR) at each detector and post-select cases where the left and right detectors each receive N photons. This forces the source photon state into a superposition of N photons arriving at either the left or right antenna. The associated phases of the photons superimpose.

The joint state output from the arms of the left and right beamsplitters will have the form

$$|\Omega_N\rangle_{LR} = \frac{1}{N!\cdot 2^N\cdot\sqrt{2}}\Big(e^{iN\phi}(\hat{c}_L^\dagger+\hat{d}_L^\dagger)^N(\hat{c}_R^\dagger-\hat{d}_R^\dagger)^N + e^{iN\delta}(\hat{c}_L^\dagger-\hat{d}_L^\dagger)^N(\hat{c}_R^\dagger+\hat{d}_R^\dagger)^N\Big)|0\rangle \tag{12}$$

$$= \frac{1}{N!\cdot 2^N\cdot\sqrt{2}}\Bigg(e^{iN\phi}\Bigg(\sum_{\ell_1=0}^{N}\binom{N}{\ell_1}(\hat{c}_L^\dagger)^{N-\ell_1}(\hat{d}_L^\dagger)^{\ell_1} \cdot\sum_{r_1=0}^{N}\binom{N}{r_1}(-1)^{r_1}(\hat{c}_R^\dagger)^{N-r_1}(\hat{d}_R^\dagger)^{r_1}\Bigg) + e^{iN\delta}\Bigg(\sum_{\ell_2=0}^{N}\binom{N}{\ell_2}(-1)^{\ell_2}(\hat{c}_L^\dagger)^{N-\ell_2}(\hat{d}_L^\dagger)^{\ell_2} \cdot\sum_{r_2=0}^{N}\binom{N}{r_2}(\hat{c}_R^\dagger)^{N-r_2}(\hat{d}_R^\dagger)^{r_2}\Bigg)\Bigg)|0\rangle \tag{13}$$

$$= \frac{1}{N!\cdot 2^N\cdot\sqrt{2}}\sum_{\ell,r=0}^{N}\Bigg[\big((-1)^r e^{iN\phi}+(-1)^\ell e^{iN\delta}\big) \cdot\binom{N}{\ell}\binom{N}{r}(\hat{c}_L^\dagger)^{N-\ell}(\hat{d}_L^\dagger)^{\ell}(\hat{c}_R^\dagger)^{N-r}(\hat{d}_R^\dagger)^{r}|0\rangle\Bigg] \tag{14}$$

$$= \frac{1}{2^N\cdot\sqrt{2}}\sum_{\ell,r=0}^{N}\Bigg[\big((-1)^r e^{iN\phi}+(-1)^\ell e^{iN\delta}\big) \cdot\sqrt{\binom{N}{\ell}}\sqrt{\binom{N}{r}}|N-\ell\rangle_{L,c}|\ell\rangle_{L,d}|N-r\rangle_{R,c}|r\rangle_{R,d}\Bigg]. \tag{15}$$

Remark 1: As mentioned before, we can adapt these calculations to account for each of the N photons at an antenna belonging to a different mode, such as a distinct time bin. In this case, in each of N consecutive time bins, we post-select cases in which we receive a single photon at each antenna. In Equation (9), the term $(\hat{a}_L^\dagger)^N$ will be replaced by $\prod_{i=1}^{N}\hat{a}_{L,i}^\dagger$, where $\hat{a}_{L,i}$ is the corresponding annihilation operator in the i^{th} mode. Also in this equation, the divisor $N!$ will be omitted. In the expression for the final output state in Equation (15), terms $\sqrt{\binom{N}{\ell}}\sqrt{\binom{N}{r}}|N-\ell\rangle_{L,c}|\ell\rangle_{L,d}|N-r\rangle_{R,c}|r\rangle_{R,d}$ will be replaced by a sum of $\binom{N}{\ell}\cdot\binom{N}{r}$ different terms corresponding to indexed modes for the ℓ and r photons in the d-outputs of the left and right antennas, respectively. It is important to note that the probabilities of ℓ and r having the same parity or opposite parities will remain the same as in Equation (15).

Note that each term of the final sum has a coefficient proportional to either $e^{iN\phi}+e^{iN\delta}$ or $e^{iN\phi}-e^{iN\delta}$, depending on whether the photon counts ℓ and r in the lower arms of the left and right beamsplitters have the same parity or opposite parities, respectively. Each of these terms corresponds to a unique signature of photon count numbers at the c and d output ports of each beamsplitter, and the probability

of the associated photon count configuration will gain a factor of $[1 + \mathrm{Re}(e^{iN(\phi-\delta)})]/2$ or $[1 - \mathrm{Re}(e^{iN(\phi-\delta)})]/2$, respectively, proportional to the coefficient's squared modulus.

Thus, by tracking the c and d output photon count numbers at each antenna, we can estimate these probability factors. If we assume that the source photons arrive from a direction given by the unit vector $\mathbf{s}$, then from Equation (4) we have

$$[1 \pm \mathrm{Re}(e^{iN(\phi-\delta)})]/2 = [1 \pm \mathrm{Re}(e^{2\pi i\nu N\mathbf{b}\cdot\mathbf{s}}e^{-iN\delta})]/2,$$

where $e^{2\pi i\nu N\mathbf{b}\cdot\mathbf{s}}$ is the contribution to the visibility $V(N\mathbf{b})$ arising from the $\mathbf{s}$ direction. Because we are considering a point source with fixed position $\mathbf{s}$, this demonstrates that we can use these probability factors to compute $V(N\mathbf{b})$ for a point source by varying the delay δ. This effectively increases our baseline distance N-fold and reduces our error in estimating the source angle θ (or the angle Φ between $\mathbf{b}$ and $\mathbf{s}$) by a factor of N by Equation (2).

IV. Estimating Point Source Location

Sharing N-photon entangled states between our two antennas is akin to using N entangled probes to perform Heisenberg-limited sensing [4, 5]. In Heisenberg-limited sensing, the factor of N increase in the parameter ϕ ultimately reduces the error (standard deviation) of our estimate from scaling as $1/\sqrt{N}$ (the classical "shot noise" limit, achievable using N independent probes without exploiting entanglement) to scaling as $1/N$ (the Heisenberg limit). A number of studies have already examined using quantum states to achieve Heisenberg-limited sensing in interferometry [6–10].

We can compute the Fisher information F of our detection scheme's ability to estimate the position of a point source, which was derived by Tsang [11] for the $N = 1$ case originally described in Gottesman et al. [3]. The Cramér-Rao bound states that the variance of our estimate of ϕ is lower-bounded by $1/F$, and this can be asymptotically achieved over long integration times by the maximum likelihood estimation. Let ϵ be the average detected source photon number. Then the probability of the N source photons and the N EPS photons arriving at opposite antennas is $(\epsilon/2)^N \frac{e^{-\epsilon}}{N!}$. Taking our measurement to be a comparison of the ℓ and r parities from Equation (15), we will denote by E_0 the event that N photons arrived at each antenna, and the ℓ and r counts have the same parity. Likewise, we let E_1 denote the event that N photons arrived at each antenna, with ℓ and r having opposite parities. We can compute $P(E_0|\phi)$ and $P(E_1|\phi)$ by summing the squared magnitudes of the corresponding coefficients in Equation (15). For instance,

$$P(E_0|\phi) = \frac{(\epsilon/2)^N \frac{e^{-\epsilon}}{N!}}{2^{2N+1}} \left|e^{iN\phi} + e^{iN\delta}\right|^2 \left[\sum_{\ell,r \text{ even}} \binom{N}{\ell}\binom{N}{r} + \sum_{\ell,r \text{ odd}} \binom{N}{\ell}\binom{N}{r}\right], \quad (16)$$

$$P(E_1|\phi) = \frac{(\epsilon/2)^N \frac{e^{-\epsilon}}{N!}}{2^{2N+1}} \left|e^{iN\phi} - e^{iN\delta}\right|^2 \left[\sum_{\substack{\ell \text{ odd},\\ r \text{ even}}} \binom{N}{\ell}\binom{N}{r} + \sum_{\substack{\ell \text{ even},\\ r \text{ odd}}} \binom{N}{\ell}\binom{N}{r}\right]. \quad (17)$$

We can substitute $\left|e^{iN\phi} + e^{iN\delta}\right|^2 = 2 \cdot \left(1 + \mathrm{Re}(e^{iN(\phi-\delta)})\right)$. Furthermore, $\sum_{\ell,r \text{ even}} \binom{N}{\ell}\binom{N}{r}$ is the product of $\sum_{\ell \text{ even}} \binom{N}{\ell}$ and $\sum_{r \text{ even}} \binom{N}{r}$, which are equal. Because $\sum_{\ell \text{ even}} \binom{N}{\ell}$ and $\sum_{\ell \text{ odd}} \binom{N}{\ell}$ are the sums of the coefficients of the even and odd exponent terms of $(x+1)^N$, respectively, we see that substituting 1 for x gives us $2^N = \sum_{\ell \text{ even}} \binom{N}{\ell} + \sum_{\ell \text{ odd}} \binom{N}{\ell}$, and substituting -1 for x gives $\sum_{\ell \text{ even}} \binom{N}{\ell} - \sum_{\ell \text{ odd}} \binom{N}{\ell} = 0$, whereby $\sum_{\ell \text{ even}} \binom{N}{\ell}$ and $\sum_{\ell \text{ odd}} \binom{N}{\ell}$ are both equal to $2^N/2$ and $\sum_{\ell,r \text{ even}} \binom{N}{\ell}\binom{N}{r} = 2^{2N}/4$. Similar computations hold for the other terms in our expressions, and it follows that

$$P(E_0|\phi) = \frac{\epsilon^N e^{-\epsilon}}{2^{N+1}N!}\left(1 + \mathrm{Re}(e^{iN(\phi-\delta)})\right), \tag{18}$$

$$P(E_1|\phi) = \frac{\epsilon^N e^{-\epsilon}}{2^{N+1}N!}\left(1 - \mathrm{Re}(e^{iN(\phi-\delta)})\right). \tag{19}$$

The Fisher information is then given by

$$F := \sum_i \frac{1}{P(E_i|\phi)} D(E_i|\phi), \tag{20}$$

$$D(E_i|\phi) := \left(\frac{\partial P(E_i|\phi)}{\partial \phi}\right)^2. \tag{21}$$

We can then compute

$$D(E_0|\phi) = \left[\frac{\partial}{\partial\phi}\left[\frac{\epsilon^N e^{-\epsilon}}{2^{N+1}N!}\left(1 + \cos(N(\phi-\delta))\right)\right]\right]^2 \tag{22}$$

$$= \frac{N^2\epsilon^{2N}e^{-2\epsilon}}{2^{2N+2}(N!)^2}\sin^2(N(\phi-\delta)), \tag{23}$$

and can verify that $D(E_1|\phi) = D(E_0|\phi)$. Subsequently,

$$F = \frac{N^2\epsilon^{2N}e^{-2\epsilon}}{2^{2N+2}(N!)^2}\sin^2(N(\phi-\delta)) \cdot \left[\frac{1}{P(E_0|\phi)} + \frac{1}{P(E_1|\phi)}\right] \tag{24}$$

$$= \frac{N^2\epsilon^{N}e^{-\epsilon}}{2^{N+1}N!}\sin^2(N(\phi-\delta)) \cdot \left[\frac{1}{1+\cos(N(\phi-\delta))} + \frac{1}{1-\cos(N(\phi-\delta))}\right] \tag{25}$$

$$= \frac{N^2\epsilon^{N}e^{-\epsilon}}{2^{N+1}N!}\sin^2(N(\phi-\delta)) \cdot \left[\frac{2}{1-\cos^2(N(\phi-\delta))}\right] \tag{26}$$

$$= \frac{N^2\epsilon^{N}e^{-\epsilon}}{2^{N}N!}. \tag{27}$$

Using Stirling's approximation, $N! \sim \sqrt{2\pi N}\,(N/e)^N$, we arrive at

$$F \sim \frac{N^{3/2}e^{-\epsilon}}{\sqrt{2\pi}}\left(\frac{\epsilon e}{2N}\right)^N. \tag{28}$$

If ϵ represents the average photon number in a single time bin, and the N photon EPSs $|\Theta_N\rangle_{LR}$ are produced by sending all N photons to the left or right antenna in the same time bin, then this expression for F represents a factor by which the Fisher information scales over time. For $\epsilon > 1$, this can give a substantial improvement over

the $N = 1$ case. If instead ϵ represents the average photon number over N time bins, and $|\Theta_N\rangle_{LR}$ represents a state in which each photon belongs to a separate time bin, then we may write $\epsilon = N\epsilon_0$, where ϵ_0 is the single-time-bin average detected source photon number. In this case, we achieve

$$F_N = \left(\frac{N^{N+2}}{N!}\right)\left(\frac{\epsilon_0}{2e^{\epsilon_0}}\right)^N \tag{29}$$

$$\sim \frac{N^{3/2}}{\sqrt{2\pi}}\left(\frac{\epsilon_0 e}{2e^{\epsilon_0}}\right)^N. \tag{30}$$

We can compare this to simply using single-photon EPSs, and repeating N times, for a total Fisher information of $NF_1 = N \cdot \frac{\epsilon_0}{2e^{\epsilon_0}}$. It can be shown that this is greater than Equation (29) for every value of N, regardless of ϵ_0. We conclude that, while we can reduce the variance of our estimate of ϕ by preparing our EPS photons for $|\Theta_N\rangle_{LR}$ all in the same time bin for a large enough average photon number, we gain no such benefit from distributing the photons over separate time bins. Furthermore, neither case allows us to outperform a linear scaling of Fisher information with respect to integration time, indicating that we cannot outperform the shot noise sensing limit. This is largely due to the fact that the probability of N source photons all arriving at a single antenna becomes diminishingly small, scaling as $O(2^{-N})$. We must build upon our method to combat this.

V. Salvaging More Events with More Robust Entangled States

We can improve our performance by considering more elaborate forms of shared entanglement between the two antennas. For instance, for N even, consider using states of the form

$$|\Psi_{N,d}\rangle^{\text{eps}} = \frac{1}{\sqrt{2}} \sum_{\substack{\mathbf{x}\in\{0,1\}^N,\\ ||\mathbf{x}||_1 = \frac{N}{2}-d}} \left(e^{i\left(\frac{N}{2}+d\right)\delta}|\mathbf{x}\rangle_L|\overline{\mathbf{x}}\rangle_R + e^{i\left(\frac{N}{2}-d\right)\delta}|\overline{\mathbf{x}}\rangle_L|\mathbf{x}\rangle_R\right). \tag{31}$$

Here, for a binary vector $\mathbf{x} \in \{0,1\}^N$, the notation $\overline{\mathbf{x}}$ represents interchanging 0 and 1 for each element of $\mathbf{x}$. In other words, over N time bins, exactly $\frac{N}{2} - d$ photons are sent to one antenna, while $\frac{N}{2} + d$ are sent to the other. We can verify that if we detect a single photon at each antenna over all the time intervals (one from the source, and one from the entangled state), our final state will be projected into the form

$$|\Theta_{N,d}\rangle_{LR} = \frac{1}{\sqrt{2}}e^{i\left(\frac{N}{2}-d\right)(\phi+\delta)} \sum_{\mathcal{S}_0} \left(e^{i2d\phi}|\mathbf{x}\rangle_L^{\text{eps}}|\mathbf{y}\rangle_L^{\text{src}}|\overline{\mathbf{x}}\rangle_R^{\text{eps}}|\overline{\mathbf{y}}\rangle_R^{\text{src}} + e^{i2d\delta}|\overline{\mathbf{x}}\rangle_L^{\text{eps}}|\overline{\mathbf{y}}\rangle_L^{\text{src}}|\mathbf{x}\rangle_R^{\text{eps}}|\mathbf{y}\rangle_R^{\text{src}}\right) \tag{32}$$

$$= \frac{1}{\sqrt{2}}e^{i\left(\frac{N}{2}-d\right)(\phi+\delta)} \left(e^{i2d\phi}\left[\sum_{\mathcal{S}_0}|\mathbf{x}\rangle_L^{\text{eps}}|\mathbf{y}\rangle_L^{\text{src}}|\overline{\mathbf{x}}\rangle_R^{\text{eps}}|\overline{\mathbf{y}}\rangle_R^{\text{src}}\right] + e^{i2d\delta}\left[\sum_{\mathcal{S}_0}|\overline{\mathbf{x}}\rangle_L^{\text{eps}}|\overline{\mathbf{y}}\rangle_L^{\text{src}}|\mathbf{x}\rangle_R^{\text{eps}}|\mathbf{y}\rangle_R^{\text{src}}\right]\right), \tag{33}$$

where

$$\mathcal{S}_0 := \{\mathbf{x}, \mathbf{y} \in \{0,1\}^N \ : \ ||\mathbf{x}||_1 = \frac{N}{2} - d, \ ||\mathbf{y}||_1 = \frac{N}{2} + d, \ \mathbf{x} + \mathbf{y} = \mathbf{1}\}$$

and $\mathbf{1}$ denotes the all-1 length-N binary vector. Paralleling our math from Equations (12)–(15), every photon count signature in the "c" and "d" outputs of the left and right antenna will correspond to one of the terms in the above sum, and the coefficient will be scaled by a factor of either $e^{i2d\phi} + e^{i2d\delta}$ or $e^{i2d\phi} - e^{i2d\delta}$. Following the same reasoning as before, by measuring photon counts, we can estimate the probabilities $[1 \pm \mathrm{Re}(e^{i2d(\phi-\delta)})]/2 = [1 \pm \mathrm{Re}(e^{2\pi i\nu(2d)\mathbf{b}\cdot\mathbf{s}}e^{-iN\delta})]/2$, as if we increased our baseline by a factor of $2d$.

The difference here is that, while we gain less of an effective baseline increase than before, we no longer require so many source photons to arrive at a single antenna. If a source photon fails to arrive, or arrives at the same antenna as one of our EPS photons, the position of the corresponding EPS photon is determined. If n_L of these events determine EPS photons arriving at the left antenna, and n_R events determine EPS photons arriving at the right, then the state of the remaining $N - n_L - n_R$ time bins is projected onto a superposition of two states, which are each described as follows: In the first superposition state (of "Type 1"), $\frac{N}{2} - d - n_L$ EPS photons arrive at the left antenna (along with $\frac{N}{2} + d - n_R$ source photons), while $\frac{N}{2} + d - n_R$ EPS photons arrive at the right antenna (along with $\frac{N}{2} - d - n_L$ source photons). In the second superposition state ("Type 2"), $\frac{N}{2} + d - n_L$ EPS photons arrive at the left antenna (as do $\frac{N}{2} - d - n_R$ source photons), and $\frac{N}{2} - d - n_R$ EPS photons arrive at the right antenna (with $\frac{N}{2} + d - n_L$ source photons). We devote the next section to proving the following theorem:

Theorem 1. By setting $d = fN$ for a fixed constant $f < \epsilon_0/4$ and using the state $|\Psi_{N,d}\rangle^{\mathrm{eps}}$ in our configuration, our detection scheme's ability to estimate the point source angle ϕ achieves a Cramér-Rao lower bound on standard deviation scaling as $\sim 1/N^{3/4}$, thereby beating the shot noise scaling limit. Furthermore, if we allow f to vary as $f = \sqrt{\frac{-\pi kN + \sqrt{\pi^2k^2N^2 + 64\pi\epsilon_0^3 kN}}{512\epsilon_0}}$ for a constant k, we achieve Heisenberg-limited scaling of $\sim 1/N$.

VI. Proof of Theorem 1

We will prove Theorem 1 in three steps. First, we will derive an expression for the output detection state arising from our setup and a corresponding expression for the Fisher information. Then we will approximate the terms arising in this expression over a typical set of (n_L, n_R). Finally, we will find a bound on the probability of these typical (n_L, n_R) combinations to obtain the scalings of the total Fisher information described in the theorem.

A. The Output Detection State

Note that every set of $n_L + n_R$ time bins is equally likely to occur. For each fixed set of n_L and n_R time bins, there are $\binom{N-n_L-n_R}{\frac{N}{2}-d-n_L}$ events of Type 1 and $\binom{N-n_L-n_R}{\frac{N}{2}+d-n_L}$ events of Type 2, each of which is equally probable. Each Type 1 state gains a phase factor of

$$e^{i[(\frac{N}{2}-d-n_L)\delta+(\frac{N}{2}+d-n_R)\phi]} = e^{i[(\frac{N}{2}-d)(\delta+\phi)-n_L\delta-n_R\phi]} \cdot e^{i2d\phi}.$$

Each Type 2 state correspondingly gains a phase factor of

$$e^{i[(\frac{N}{2}+d-n_L)\delta+(\frac{N}{2}-d-n_R)\phi]} = e^{i[(\frac{N}{2}-d)(\delta+\phi)-n_L\delta-n_R\phi]} \cdot e^{i2d\delta}.$$

The result is a state in the form

$$|\Theta_{N,d,n_L,n_R}\rangle_{LR} = \frac{e^{i[(\frac{N}{2}-d)(\delta+\phi)-n_L\delta-n_R\phi]}}{\sqrt{\binom{N-n_L-n_R}{\frac{N}{2}-d-n_L}+\binom{N-n_L-n_R}{\frac{N}{2}+d-n_L}}} \times \left(\sum_{|T_1^{(\mathbf{x},\mathbf{y})}\rangle \in \mathcal{T}_1^{(n_L,n_R)}} e^{i2d\phi}|T_1^{(\mathbf{x},\mathbf{y})}\rangle + \sum_{|T_2^{(\mathbf{x},\mathbf{y})}\rangle \in \mathcal{T}_2^{(n_L,n_R)}} e^{i2d\delta}|T_2^{(\mathbf{x},\mathbf{y})}\rangle \right), \tag{34}$$

where $\mathcal{T}_1^{(n_L,n_R)}$ and $\mathcal{T}_2^{(n_L,n_R)}$ are the Type 1 and Type 2 states, respectively:

$$\mathcal{T}_1^{(n_L,n_R)} := \left\{ |\mathbf{x}\rangle_L^{\text{eps}}|\mathbf{y}\rangle_L^{\text{src}}|\overline{\mathbf{x}}\rangle_R^{\text{eps}}|\overline{\mathbf{y}}\rangle_R^{\text{src}} \,\middle|\, \mathbf{x},\mathbf{y} \in \{0,1\}^{N-n_L-n_R}, \; ||\mathbf{x}||_1 = \frac{N}{2} - d - n_L, \; \mathbf{x}+\mathbf{y} = \mathbf{1} \right\}, \tag{35}$$

$$\mathcal{T}_2^{(n_L,n_R)} := \left\{ |\mathbf{x}\rangle_L^{\text{eps}}|\mathbf{y}\rangle_L^{\text{src}}|\overline{\mathbf{x}}\rangle_R^{\text{eps}}|\overline{\mathbf{y}}\rangle_R^{\text{src}} \,\middle|\, \mathbf{x},\mathbf{y} \in \{0,1\}^{N-n_L-n_R}, \; ||\mathbf{x}||_1 = \frac{N}{2} + d - n_L, \; \mathbf{x}+\mathbf{y} = \mathbf{1} \right\}, \tag{36}$$

and $|T_1^{(\mathbf{x},\mathbf{y})}\rangle$ and $|T_2^{(\mathbf{x},\mathbf{y})}\rangle$ represent the corresponding elements of $\mathcal{T}_1^{(n_L,n_R)}$ and $\mathcal{T}_2^{(n_L,n_R)}$. Notice that the vector $\mathbf{y}$ implicitly satisfies $||\mathbf{y}||_1 = \frac{N}{2} + d - n_R$ in $\mathcal{T}_1^{(n_L,n_R)}$ and $||\mathbf{y}||_1 = \frac{N}{2} - d - n_R$ in $\mathcal{T}_2^{(n_L,n_R)}$.

We can reexpress Equation (34) as

$$|\Theta_{N,d,n_L,n_R}\rangle_{LR} = e^{i[(\frac{N}{2}-d)(\delta+\phi)-n_L\delta-n_R\phi]} \cdot \left(e^{i2d\phi}\sqrt{A_1^{(n_L,n_R)}}|\mathbf{T}_1^{(n_L,n_R)}\rangle + e^{i2d\delta}\sqrt{A_2^{(n_L,n_R)}}|\mathbf{T}_2^{(n_L,n_R)}\rangle \right), \tag{37}$$

where

$$|\mathbf{T}_1^{(n_L,n_R)}\rangle := \frac{1}{\sqrt{\binom{N-n_L-n_R}{\frac{N}{2}-d-n_L}}} \sum_{|T_1^{(\mathbf{x},\mathbf{y})}\rangle \in \mathcal{T}_1^{(n_L,n_R)}} |T_1^{(\mathbf{x},\mathbf{y})}\rangle,$$

$$|\mathbf{T}_2^{(n_L,n_R)}\rangle := \frac{1}{\sqrt{\binom{N-n_L-n_R}{\frac{N}{2}+d-n_L}}} \sum_{|T_2^{(\mathbf{x},\mathbf{y})}\rangle \in \mathcal{T}_2^{(n_L,n_R)}} |T_2^{(\mathbf{x},\mathbf{y})}\rangle,$$

$$A_1^{(n_L,n_R)} := \frac{\binom{N-n_L-n_R}{\frac{N}{2}-d-n_L}}{\binom{N-n_L-n_R}{\frac{N}{2}-d-n_L}+\binom{N-n_L-n_R}{\frac{N}{2}+d-n_L}} = \frac{\prod_{k=-d}^{d}\left(\frac{N}{2}+k-n_L\right)}{\prod_{k=-d}^{d}\left(\frac{N}{2}+k-n_L\right)+\prod_{k=-d}^{d}\left(\frac{N}{2}+k-n_R\right)}, \tag{38}$$

$$A_2^{(n_L,n_R)} := \frac{\binom{N-n_L-n_R}{\frac{N}{2}+d-n_L}}{\binom{N-n_L-n_R}{\frac{N}{2}+d-n_L}+\binom{N-n_L-n_R}{\frac{N}{2}-d-n_L}} = \frac{\prod_{k=-d}^{d}\left(\frac{N}{2}+k-n_R\right)}{\prod_{k=-d}^{d}\left(\frac{N}{2}+k-n_L\right)+\prod_{k=-d}^{d}\left(\frac{N}{2}+k-n_R\right)}, \tag{39}$$

where we note that $A_1^{(n_L,n_R)} + A_2^{(n_L,n_R)} = 1$. We can map each combination of states $|T_1^{(\mathbf{x},\mathbf{y})}\rangle \in \mathcal{T}_1^{(n_L,n_R)}$ and $|T_2^{(\mathbf{x},\mathbf{y})}\rangle \in \mathcal{T}_2^{(n_L,n_R)}$ to a state of the photon count signatures in the "c" and "d" output ports of the left and right antennas using a process similar to that of Equations (12)–(15). For any (n_L, n_R) pair, we can index the output events by the vectors

$$\mathcal{Z}^{(n_L,n_R)} := \{(\mathbf{z}_L, \mathbf{z}_R) \in \{0,1\}^{N-n_L-n_R} \times \{0,1\}^{N-n_L-n_R}\}, \tag{40}$$

where $\mathbf{z}_L$ contains a 1 in the i^{th} position if a detection event occurred at the "d" output port of the left antenna at time index i out of the $N - n_L - n_R$ timeslots, and $\mathbf{z}_R$ is defined similarly. A given state $|T_1^{(\mathbf{x},\mathbf{y})}\rangle$ will map to a linear combination of the output states $|(\mathbf{z}_L, \mathbf{z}_R)\rangle$, each with a coefficient gaining a factor of -1 for every index where $\mathbf{z}_L$ and $\mathbf{x}$ are both nonzero, and for every index where $\mathbf{z}_R$ and $\overline{\mathbf{x}}$ are both nonzero. The same holds for each state $|T_2^{(\mathbf{x},\mathbf{y})}\rangle$. We can write this concisely as

$$|T_1^{(\mathbf{x},\mathbf{y})}\rangle \mapsto \frac{1}{\sqrt{2^{2(N-n_L-n_R)}}} \sum_{\mathcal{Z}^{(n_L,n_R)}} (-1)^{\mathbf{z}_L\cdot\mathbf{x}+\mathbf{z}_R\cdot\overline{\mathbf{x}}}|(\mathbf{z}_L, \mathbf{z}_R)\rangle \tag{41}$$

$$= \frac{1}{2^{N-n_L-n_R}} \sum_{\mathcal{Z}^{(n_L,n_R)}} (-1)^{\mathbf{z}_L\cdot\mathbf{x}+\mathbf{z}_R\cdot(\mathbf{1}-\mathbf{x})}|(\mathbf{z}_L, \mathbf{z}_R)\rangle \tag{42}$$

$$= \frac{1}{2^{N-n_L-n_R}} \sum_{\mathcal{Z}^{(n_L,n_R)}} (-1)^{\mathbf{z}_R\cdot\mathbf{1}} \cdot (-1)^{(\mathbf{z}_L-\mathbf{z}_R)\cdot\mathbf{x}}|(\mathbf{z}_L, \mathbf{z}_R)\rangle, \tag{43}$$

where $\mathbf{1}$ is the vector of all ones. Similarly,

$$|T_2^{(\mathbf{x},\mathbf{y})}\rangle \mapsto \frac{1}{2^{N-n_L-n_R}} \sum_{\mathcal{Z}^{(n_L,n_R)}} (-1)^{\mathbf{z}_R\cdot\mathbf{1}} \cdot (-1)^{(\mathbf{z}_L-\mathbf{z}_R)\cdot\mathbf{x}}|(\mathbf{z}_L, \mathbf{z}_R)\rangle. \tag{44}$$

It follows that the states $|\mathbf{T}_1^{(n_L,n_R)}\rangle$ and $|\mathbf{T}_2^{(n_L,n_R)}\rangle$ will map to linear combinations of output states of the forms

$$|\mathbf{T}_1^{(n_L,n_R)}\rangle \mapsto \frac{2^{-(N-n_L-n_R)}}{\sqrt{\binom{N-n_L-n_R}{\frac{N}{2}-d-n_L}}} \sum_{\mathcal{Z}^{(n_L,n_R)}} (-1)^{\mathbf{z}_R\cdot\mathbf{1}} \left(\sum_{||\mathbf{x}||_1=\frac{N}{2}-d-n_L} (-1)^{(\mathbf{z}_L-\mathbf{z}_R)\cdot\mathbf{x}} \right) |(\mathbf{z}_L,\mathbf{z}_R)\rangle \tag{45}$$

$$= 2^{-(N-n_L-n_R)} \sum_{\mathcal{Z}^{(n_L,n_R)}} |Z_1^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)\rangle \tag{46}$$

$$=: |\mathbf{Z}_1^{(n_L,n_R)}\rangle, \tag{47}$$

$$|\mathbf{T}_2^{(n_L,n_R)}\rangle \mapsto \frac{2^{-(N-n_L-n_R)}}{\sqrt{\binom{N-n_L-n_R}{\frac{N}{2}+d-n_L}}} \sum_{\mathcal{Z}^{(n_L,n_R)}} (-1)^{\mathbf{z}_R\cdot\mathbf{1}} \left(\sum_{||\mathbf{x}||_1=\frac{N}{2}+d-n_L} (-1)^{(\mathbf{z}_L-\mathbf{z}_R)\cdot\mathbf{x}} \right) |(\mathbf{z}_L,\mathbf{z}_R)\rangle \tag{48}$$

$$= 2^{-(N-n_L-n_R)} \sum_{\mathcal{Z}^{(n_L,n_R)}} |Z_2^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)\rangle \tag{49}$$

$$=: |\mathbf{Z}_2^{(n_L,n_R)}\rangle, \tag{50}$$

where we define

$$|Z_1^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)\rangle := Z_1^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)\cdot|(\mathbf{z}_L,\mathbf{z}_R)\rangle, \tag{51}$$

$$|Z_2^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)\rangle := Z_2^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)\cdot|(\mathbf{z}_L,\mathbf{z}_R)\rangle, \tag{52}$$

with complex coefficients $Z_1^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)$ and $Z_2^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)$ defined by

$$Z_1^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R) := \frac{1}{\sqrt{\binom{N-n_L-n_R}{\frac{N}{2}-d-n_L}}}(-1)^{\mathbf{z}_R\cdot\mathbf{1}} \left(\sum_{||\mathbf{x}||_1=\frac{N}{2}-d-n_L} (-1)^{(\mathbf{z}_L-\mathbf{z}_R)\cdot\mathbf{x}} \right), \tag{53}$$

$$Z_2^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R) := \frac{1}{\sqrt{\binom{N-n_L-n_R}{\frac{N}{2}+d-n_L}}}(-1)^{\mathbf{z}_R\cdot\mathbf{1}} \left(\sum_{||\mathbf{x}||_1=\frac{N}{2}+d-n_L} (-1)^{(\mathbf{z}_L-\mathbf{z}_R)\cdot\mathbf{x}} \right). \tag{54}$$

We can compute $Z_1^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)$ and $Z_2^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)$ using the following lemma:

Lemma 1. The sum $\sum_{\mathbf{x}\in\mathcal{S}_k}(-1)^{\mathbf{w}\cdot\mathbf{x}}$, where $\mathcal{S}_k = \{\mathbf{x}\in\{0,1\}^m \ : \ ||\mathbf{x}||_1 = k\}$ and $\mathbf{w}\in\{0,1\}^m$ has weight $||\mathbf{w}||_1 = \ell$, is equal to the coefficient of x^k in the polynomial expansion of $(1-x)^\ell(1+x)^{m-\ell}$.

Proof. We can write $(-1)^{\mathbf{w}\cdot\mathbf{x}} = (-1)^{|\mathrm{supp}(\mathbf{w})\cap\mathrm{supp}(\mathbf{x})|}$. We can then separate the vectors $\mathbf{x}$ by the size of the overlap of their support with that of $\mathbf{w}$: For each value of $s \le \min(k,\ell)$, there are $\binom{\ell}{s}$ ways to choose an overlap of size s and $\binom{m-\ell}{k-s}$ ways of choosing the rest of the support of $\mathbf{x}$. Thus,

$$\sum_{\mathbf{x}\in\mathcal{S}_k}(-1)^{\mathbf{w}\cdot\mathbf{x}} = \sum_{s=0}^{\min(k,\ell)}(-1)^s\binom{\ell}{s}\binom{m-\ell}{k-s}. \tag{55}$$

Now, examining the coefficients of $(1-x)^\ell(1+x)^{m-\ell}$,

$$(1-x)^\ell(1+x)^{m-\ell} = \left(\sum_{i_1=0}^{\ell}(-1)^{i_1}\binom{\ell}{i_1}x^{i_1}\right)\left(\sum_{i_2=0}^{m-\ell}\binom{m-\ell}{i_2}x^{i_2}\right) \tag{56}$$

$$= \sum_{r=0}^{m}\left(\sum_{i_3=0}^{r}(-1)^{i_3}\binom{\ell}{i_3}\binom{m-\ell}{r-i_3}\right)x^r. \tag{57}$$

We see that the coefficient of x^k is identical to our above calculation. □

Note that from the definitions in Equations (35) and (36), we may write

$$\sum_{||\mathbf{x}||_1=\frac{N}{2}-d-n_L}(-1)^{(\mathbf{z}_L-\mathbf{z}_R)\cdot\mathbf{x}} = (-1)^{-(\mathbf{z}_L-\mathbf{z}_R)\cdot\mathbf{1}}\sum_{||\mathbf{x}||_1=\frac{N}{2}-d-n_L}(-1)^{(\mathbf{z}_L-\mathbf{z}_R)\cdot(\mathbf{1}+\mathbf{x})} \tag{58}$$

$$= (-1)^{(\mathbf{z}_L-\mathbf{z}_R)\cdot\mathbf{1}}\sum_{||\mathbf{y}||_1=\frac{N}{2}+d-n_R}(-1)^{(\mathbf{z}_L-\mathbf{z}_R)\cdot\mathbf{y}}. \tag{59}$$

This, along with the fact that $\binom{N-n_L-n_R}{\frac{N}{2}-d-n_R} = \binom{N-n_L-n_R}{\frac{N}{2}+d-n_L}$, gives us the relation

$$Z_1^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R) = (-1)^{(\mathbf{z}_L-\mathbf{z}_R)\cdot\mathbf{1}}\cdot Z_2^{(n_R,n_L)}(\mathbf{z}_L,\mathbf{z}_R), \tag{60}$$

which represents the fact that the coefficients of $|Z_1^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)\rangle$ in $|\mathbf{Z}_1^{(n_L,n_R)}\rangle$ and $|Z_2^{(n_R,n_L)}(\mathbf{z}_L,\mathbf{z}_R)\rangle$ in $|\mathbf{Z}_2^{(n_R,n_L)}\rangle$ differ by a factor of $(-1)^{(\mathbf{z}_L-\mathbf{z}_R)\cdot\mathbf{1}}$. With a similar reasoning applied to Equations (38) and (39), we have

$$A_1^{(n_L,n_R)} = A_2^{(n_R,n_L)}. \tag{61}$$

We can express the output detection state resulting from the received state $|\Theta_{N,d,n_L,n_R}\rangle_{LR}$ from Equation (37) as

$$|\Theta_{N,d,n_L,n_R}\rangle_{LR} \mapsto |\Phi_{N,d,n_L,n_R}\rangle_{LR}$$

$$= e^{i[(\frac{N}{2}-d)(\delta+\phi)-n_L\delta-n_R\phi]}\left(e^{i2d\phi}\sqrt{A_1^{(n_L,n_R)}}|\mathbf{Z}_1^{(n_L,n_R)}\rangle + e^{i2d\delta}\sqrt{A_2^{(n_L,n_R)}}|\mathbf{Z}_2^{(n_L,n_R)}\rangle\right) \tag{62}$$

$$= \frac{e^{i[(\frac{N}{2}-d)(\delta+\phi)-n_L\delta-n_R\phi]}}{2^{(N-n_L-n_R)}}\sum_{\mathcal{Z}^{(n_L,n_R)}}\left(e^{i2d\phi}\sqrt{A_1^{(n_L,n_R)}}|Z_1^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)\rangle + e^{i2d\delta}\sqrt{A_2^{(n_L,n_R)}}|Z_2^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)\rangle\right) \tag{63}$$

$$= \frac{e^{i[(\frac{N}{2}-d)(\delta+\phi)-n_L\delta-n_R\phi]}}{2^{(N-n_L-n_R)}}\sum_{\mathcal{Z}^{(n_L,n_R)}}\left[\left(e^{i2d\phi}\sqrt{A_1^{(n_L,n_R)}}\cdot Z_1^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R) + e^{i2d\delta}\sqrt{A_2^{(n_L,n_R)}}\cdot Z_2^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)\right)|(\mathbf{z}_L,\mathbf{z}_R)\rangle\right]. \tag{64}$$

Every detection event corresponds to a specific instance of the combination $\{n_L, n_R, (\mathbf{z}_L, \mathbf{z}_R)\}$. The probability of such an event depends on $P(n_L, n_R)$, the probability of observing n_L and n_R, and the squared norm of the coefficient of $|(\mathbf{z}_L, \mathbf{z}_R)\rangle$ in Equation (64):

$$P(\{n_L, n_R, (\mathbf{z}_L, \mathbf{z}_R)\}) = \frac{P(n_L, n_R)}{2^{2(N-n_L-n_R)}} \cdot \left| \left(e^{i2d\phi} \sqrt{A_1^{(n_L,n_R)}} \cdot Z_1^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R) + e^{i2d\delta} \sqrt{A_2^{(n_L,n_R)}} \cdot Z_2^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R) \right) \right|^2 \quad (65)$$

$$= \frac{P(n_L, n_R)}{2^{2(N-n_L-n_R)}} \cdot \left[A_1^{(n_L,n_R)} \cdot \left(Z_1^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R) \right)^2 + A_2^{(n_L,n_R)} \cdot \left(Z_2^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R) \right)^2 + 2\sqrt{A_1^{(n_L,n_R)} \cdot A_2^{(n_L,n_R)}} \cdot Z_1^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R) \cdot Z_2^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R) \cos(2d(\phi - \delta)) \right]. \quad (66)$$

Ultimately, we are interested in the contribution of each of these terms to the Fisher information. To this end, we compute

$$\left(\frac{\partial P(\{n_L, n_R, (\mathbf{z}_L, \mathbf{z}_R)\})}{\partial \phi} \right)^2 = \left(\frac{P(n_L, n_R)}{2^{2(N-n_L-n_R)}} \right)^2 \cdot 16d^2 \cdot \sin^2(2d(\phi - \delta)) \times \left(\sqrt{A_1^{(n_L,n_R)} \cdot A_2^{(n_L,n_R)}} \cdot Z_1^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R) \cdot Z_2^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R) \right)^2. \quad (67)$$

Each of these terms will give a contribution to the Fisher information of

$$F(\{n_L, n_R, (\mathbf{z}_L, \mathbf{z}_R)\}) := \frac{1}{P(\{n_L, n_R, (\mathbf{z}_L, \mathbf{z}_R)\})} \cdot \left(\frac{\partial P(\{n_L, n_R, (\mathbf{z}_L, \mathbf{z}_R)\})}{\partial \phi} \right)^2. \quad (68)$$

The total Fisher information will then be the sum

$$F = \sum_{\{n_L, n_R, (\mathbf{z}_L, \mathbf{z}_R)\}} F(\{n_L, n_R, (\mathbf{z}_L, \mathbf{z}_R)\}). \quad (69)$$

B. Approximating $A_1^{(n_L,n_R)}$ and $A_2^{(n_L,n_R)}$ for Typical (n_L, n_R)

Now, let us consider only events with

$$\frac{1}{2}\left(1 - \frac{\epsilon_0}{2}\right) N - \beta N \leq n_L \leq \frac{1}{2}\left(1 - \frac{\epsilon_0}{2}\right) N + \beta N, \quad (70)$$

$$\frac{1}{2}\left(1 - \frac{\epsilon_0}{2}\right) N - \beta N \leq n_R \leq \frac{1}{2}\left(1 - \frac{\epsilon_0}{2}\right) N + \beta N, \quad (71)$$

where ϵ_0 is the average photon number detected within a time bin, assumed to be very small, and β is a small fixed constant. Our reasoning is that in any time bin, a source photon has a probability of approximately $\epsilon_0/2$ of arriving at the opposite antenna of the corresponding EPS photon. Conversely, there is a probability $(1 - \frac{\epsilon_0}{2})$ of the photon not arriving, or arriving at the same antenna as the EPS photon, in either case

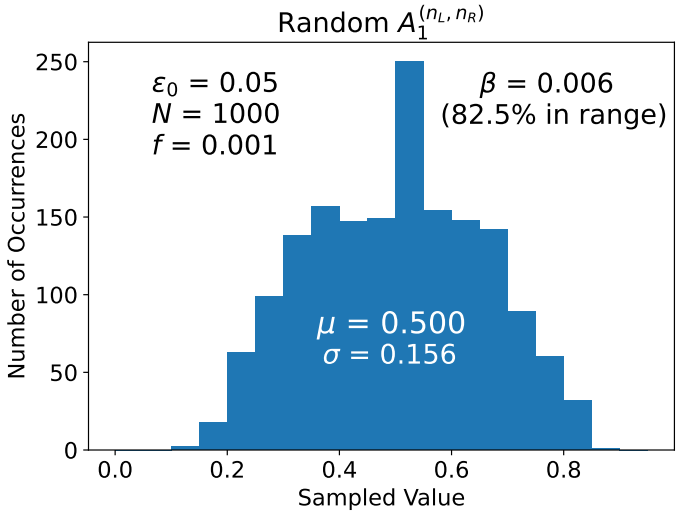


Figure 4. Simulated values of $A_1^{(n_L,n_R)}$, showing that most of the simulations yield a value of $A_1^{(n_L,n_R)}$ close to $\frac{1}{2}$. Consequently, the same will be true of $A_2^{(n_L,n_R)}$.

determining the position of the EPS photon. Therefore, the expected number of these types of time bins is $\left(1 - \frac{\epsilon_0}{2}\right) N$, half of which should correspond to n_L and half to n_R, so the above ranges should capture the detection event signatures with high probability. Furthermore, taking $d = fN$ for a small constant f, then for n_L and n_R in the above ranges, we can approximate the expressions in Equations (38) and (39) as

$$A_1^{(n_L,n_R)} \approx \frac{\left(\frac{\epsilon_0 N}{4}\right)^{2fN}}{\left(\frac{\epsilon_0 N}{4}\right)^{2fN} + \left(\frac{\epsilon_0 N}{4}\right)^{2fN}} = \frac{1}{2}, \tag{72}$$

$$A_2^{(n_L,n_R)} \approx \frac{\left(\frac{\epsilon_0 N}{4}\right)^{2fN}}{\left(\frac{\epsilon_0 N}{4}\right)^{2fN} + \left(\frac{\epsilon_0 N}{4}\right)^{2fN}} = \frac{1}{2}. \tag{73}$$

This allows us to simplify our Equations (66) and (67) as

$$\begin{aligned} P(\{n_L, n_R, (\mathbf{z}_L, \mathbf{z}_R)\}) \approx \frac{P(n_L, n_R)}{2 \cdot 2^{2(N-n_L-n_R)}} \\ \times \Big[\left(Z_1^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R) \right)^2 + \left(Z_2^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R) \right)^2 \\ + 2 \cdot Z_1^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R) \cdot Z_2^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R) \cos(2d(\phi - \delta)) \Big], \end{aligned} \tag{74}$$

$$\begin{aligned} \left(\frac{\partial P(\{n_L, n_R, (\mathbf{z}_L, \mathbf{z}_R)\})}{\partial \phi} \right)^2 \approx \left(\frac{P(n_L, n_R)}{2^{2(N-n_L-n_R)}} \right)^2 \cdot 8d^2 \cdot \sin^2(2d(\phi - \delta)) \\ \times \left(Z_1^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R) \cdot Z_2^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R) \right)^2. \end{aligned} \tag{75}$$

In Figure 4, we show the results of simulating 2,000 detection signatures and computing the value of $A_1^{(n_L,n_R)}$ for $\epsilon_0 = 0.05$, $N = 1000$, and $f = 0.001$ ($d = 1$).

Setting $\beta = 0.006$, we find that 82.5% of the simulated events lie in the ranges determined by Equations (70) and (71). As we can see, most of the simulations yield a value of $A_1^{(n_L,n_R)}$ close to $\frac{1}{2}$, justifying the approximation in Equation (72) (and consequently Equation (73), because $A_1^{(n_L,n_R)}$ and $A_2^{(n_L,n_R)}$ sum to 1).

C. Approximating $Z_1^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R)$ and $Z_2^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R)$

Now we further restrict our attention to the vectors $(\mathbf{z}_L, \mathbf{z}_R) \in \mathcal{Z}^{(n_L,n_R)}$ satisfying

$$\frac{\epsilon_0 N}{4} - \gamma N \leq ||\mathbf{z}_L - \mathbf{z}_R||_1 \leq \frac{\epsilon_0 N}{4} + \gamma N \tag{76}$$

for some small constant $\gamma < \frac{\epsilon_0}{4}$.

We reason that this will comprise a majority of the vectors $(\mathbf{z}_L, \mathbf{z}_R)$ because our above conditions on n_L and n_R imply that

$$\frac{\epsilon_0 N}{2} - \beta N \leq N - n_L - n_R \leq \frac{\epsilon_0 N}{2} + \beta N, \tag{77}$$

giving us that $N - n_L - n_R \approx \frac{\epsilon_0 N}{2}$, which will be the approximate dimension of both $\mathbf{z}_L$ and $\mathbf{z}_R$. If we choose these vectors uniformly at random, their entries will disagree in approximately half of these coordinates, the number of which will be $||\mathbf{z}_L - \mathbf{z}_R||_1$. In fact, because a given entry of $\mathbf{z}_L$ will agree with the corresponding entry of $\mathbf{z}_R$ with probability $\frac{1}{2}$, then if we define

$$\Delta^{(n_L,n_R)} := \min\left(\left|\frac{1}{2}(N - n_L - n_R) - \frac{\epsilon_0 N}{4} - \gamma N\right|, \left|\frac{1}{2}(N - n_L - n_R) - \frac{\epsilon_0 N}{4} + \gamma N\right|\right) \tag{78}$$

Hoeffding's inequality [12] gives us that

$$P\left(\frac{\epsilon_0 N}{4} - \gamma N \leq ||\mathbf{z}_L - \mathbf{z}_R||_1 \leq \frac{\epsilon_0 N}{4} + \gamma N \,\middle|\, n_L, n_R\right)$$

$$\geq 1 - 2\exp\left(-2(N - n_L - n_R)\left(\frac{\Delta^{(n_L,n_R)}}{N - n_L - n_R}\right)^2\right) \tag{79}$$

$$= 1 - 2\exp\left(-\frac{2(\Delta^{(n_L,n_R)})^2}{N - n_L - n_R}\right). \tag{80}$$

Furthermore, from Equation (77), we can see that

$$\left(\Delta^{(n_L,n_R)}\right)^2 \geq \left(\frac{\beta}{2} - \gamma\right)^2 \cdot N^2 \tag{81}$$

and

$$P\left(\frac{\epsilon_0 N}{4} - \gamma N \leq ||\mathbf{z}_L - \mathbf{z}_R||_1 \leq \frac{\epsilon_0 N}{4} + \gamma N \,\middle|\, n_L, n_R\right)$$

$$\geq 1 - 2\exp\left(-2\frac{\left(\frac{\beta}{2} - \gamma\right)^2}{\frac{\epsilon_0}{2} + \beta}N\right) \tag{82}$$

$$= 1 - 2\exp\left(-\frac{(\beta - 2\gamma)^2}{\epsilon_0 + 2\beta}N\right) \tag{83}$$

which approaches 1 as N grows large as long as $\beta \neq 2\gamma$. This allows us to bound the total number of $(\mathbf{z}_L, \mathbf{z}_R)$ satisfying Equation (76) as

$$\#\{(\mathbf{z}_L, \mathbf{z}_R) \in \mathcal{Z}^{(n_L,n_R)} \mid \frac{\epsilon_0 N}{2} - \beta N \leq N - n_L - n_R \leq \frac{\epsilon_0 N}{2} + \beta N\} \\ \geq 2^{2(N-n_L-n_R)} \cdot \left(1 - 2\exp\left(-\frac{(\beta - 2\gamma)^2}{\epsilon_0 + 2\beta} N\right)\right). \tag{84}$$

For the values of n_L, n_R, and $(\mathbf{z}_L, \mathbf{z}_R)$ considered, we can apply Lemma 1 to (53) and (54) to approximate the values of $Z_1^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R)$ and $Z_2^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R)$. In particular, we can write

$$\left(\frac{\epsilon_0}{4} - f - \beta\right) N \leq \frac{N}{2} - d - n_L \leq \left(\frac{\epsilon_0}{4} - f + \beta\right) N, \tag{85}$$

$$\left(\frac{\epsilon_0}{4} + f - \beta\right) N \leq \frac{N}{2} + d - n_L \leq \left(\frac{\epsilon_0}{4} + f + \beta\right) N. \tag{86}$$

Now, considering $Z_1^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R)$ in the language of Lemma 1, we have

$$\mathbf{w} := \mathbf{z}_L - \mathbf{z}_R, \tag{87}$$

$$m := \dim(\mathbf{w}) = N - n_L - n_R \approx \frac{\epsilon_0 N}{2}, \tag{88}$$

$$\ell := ||\mathbf{w}||_1 \approx \frac{\epsilon_0 N}{4}, \tag{89}$$

$$k := \frac{N}{2} - d - n_L \approx \left(\frac{\epsilon_0}{4} - f\right) N, \tag{90}$$

and we see that the sum $\sum_{||\mathbf{x}||_1 = \frac{N}{2} - d - n_L} (-1)^{(\mathbf{z}_L - \mathbf{z}_R)\cdot \mathbf{x}}$ is approximately equal to the coefficient of $x^{\left(\frac{\epsilon_0}{4} - f\right)N}$ in the expansion of $(1-x)^{\frac{\epsilon_0 N}{4}} \cdot (1+x)^{\frac{\epsilon_0 N}{4}} = (1-x^2)^{\frac{\epsilon_0 N}{4}}$, which itself is the coefficient of $x^{\left(\frac{\epsilon_0}{8} - \frac{f}{2}\right)N}$ in the expansion of $(1-x)^{\frac{\epsilon_0 N}{4}}$. This gives us

$$\left| \sum_{||\mathbf{x}||_1 = \frac{N}{2} - d - n_L} (-1)^{(\mathbf{z}_L - \mathbf{z}_R)\cdot \mathbf{x}} \right| \approx \binom{\frac{\epsilon_0 N}{4}}{\frac{\epsilon_0 N}{8} - \frac{fN}{2}}, \tag{91}$$

and from Stirling's approximation, we have

$$\left| \sum_{||\mathbf{x}||_1 = \frac{N}{2} - d - n_L} (-1)^{(\mathbf{z}_L - \mathbf{z}_R)\cdot \mathbf{x}} \right| \approx \sqrt{\frac{8\epsilon_0}{\pi(\epsilon_0^2 - 16f^2)N}} \\ \times \left[\left(\frac{\epsilon_0}{2\epsilon_0 - 2f}\right)^{\frac{\epsilon_0}{8} - \frac{f}{2}} \cdot \left(\frac{\epsilon_0}{2\epsilon_0 + 2f}\right)^{\frac{\epsilon_0}{8} + \frac{f}{2}} \right]^N. \tag{92}$$

Note that for Equation (91) and our other approximations to hold, we must require that

$$f < \frac{\epsilon_0}{4}. \tag{93}$$

We can use the same techniques to approximate

$$\binom{N-n_L-n_R}{\frac{N}{2}-d-n_L}^{-1/2} \approx \binom{\frac{\epsilon_0 N}{2}}{\frac{\epsilon_0 N}{4}-fN}^{-1/2} \tag{94}$$

$$\approx \left(\frac{4\epsilon_0}{\pi(\epsilon_0^2-16f^2)N}\right)^{-1/4} \cdot \left[\left(\frac{\epsilon_0}{2\epsilon_0-2f}\right)^{\frac{\epsilon_0}{4}-f} \cdot \left(\frac{\epsilon_0}{2\epsilon_0+2f}\right)^{\frac{\epsilon_0}{4}+f}\right]^{-N/2} \tag{95}$$

$$= 2^{1/4} \cdot \left(\frac{8\epsilon_0}{\pi(\epsilon_0^2-16f^2)N}\right)^{-1/4} \cdot \left[\left(\frac{\epsilon_0}{2\epsilon_0-2f}\right)^{\frac{\epsilon_0}{8}-\frac{f}{2}} \left(\frac{\epsilon_0}{2\epsilon_0+2f}\right)^{\frac{\epsilon_0}{8}+\frac{f}{2}}\right]^{-N}. \tag{96}$$

Combining all of these, Equation (53) gives us

$$\left|Z_1^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)\right| \approx \left(\frac{16\epsilon_0}{\pi(\epsilon_0^2-16f^2)N}\right)^{1/4}. \tag{97}$$

We can follow a similar line of analysis for $Z_2^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)$, using $k := \frac{N}{2}+d-n_L \approx \left(\frac{\epsilon_0}{4}+f\right)N$ in Lemma 1. Taking note of the symmetry between $+f$ and $-f$ in our approximations, and referring to Equation (54), we again find that

$$\left|Z_2^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)\right| \approx \left(\frac{16\epsilon_0}{\pi(\epsilon_0^2-16f^2)N}\right)^{1/4}. \tag{98}$$

Whether $Z_1^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)$ and $Z_2^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)$ have the same sign or opposite signs depends on the specific values of N, n_L, n_R, and d.

Remark 2: In practice, we often find that one of three scenarios tends to occur: $Z_1^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)$ and $Z_2^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)$ either 1) almost always have the same sign, 2) almost always have opposite signs, or 3) have similar and opposite signs with approximately the same frequency when sampling over (n_L,n_R) and $(\mathbf{z}_L,\mathbf{z}_R)$. In Figure 5, we demonstrate this via simulation in two cases: $(N,d)=(300,1)$ and $(400,2)$, sampling 2,000 events for each with $\epsilon_0=0.05$. In the first configuration, the values of $Z_1^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)$ and $Z_2^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)$ have positive or negative signs with roughly the same frequency, but their ratio is predominantly ≈ -1. In the second configuration, while the values of $Z_1^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)$ and $Z_2^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)$ are mostly positive, there are enough negative values that their ratio is bimodally distributed around both -1 and 1.

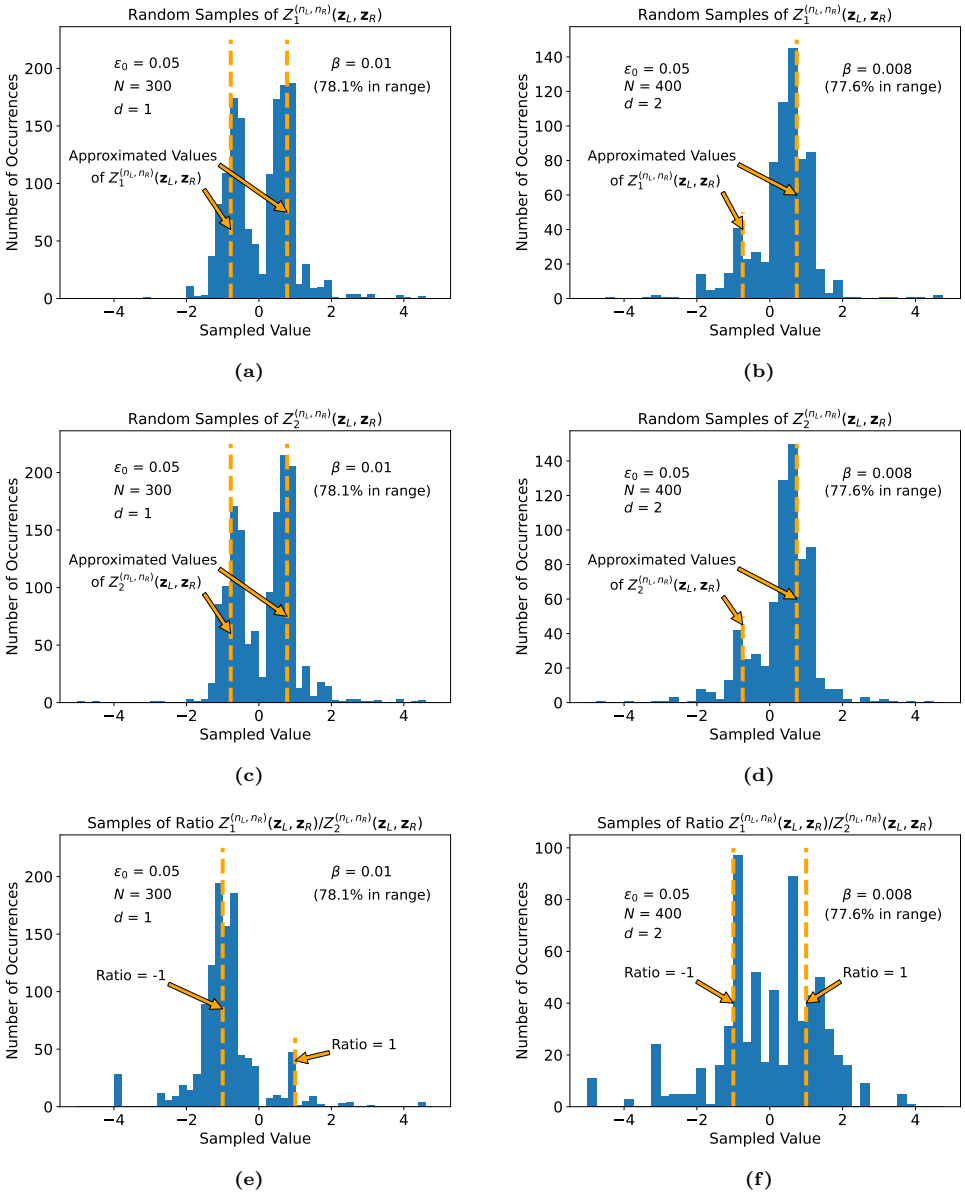


Figure 5. Simulated values of $Z_1^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)$ (top row), $Z_2^{(n_L,n_R)}(\mathbf{z}_L,\mathbf{z}_R)$ (middle row), and their ratio (bottom row) for two configurations: the left column uses parameters $N = 300$, $d = 1$, and $\beta = 0.01$, while the right column uses $N = 400$, $d = 2$, and $\beta = 0.008$. The values tend to cluster around the approximations given in Equations (97) and (98), indicated by the dotted orange lines.

We can now simplify our Equations (74) and (75) to

$$P(\{n_L, n_R, (\mathbf{z}_L, \mathbf{z}_R)\}) \approx \frac{P(n_L, n_R)}{2^{2(N-n_L-n_R)}} \cdot \sqrt{\frac{16\epsilon_0}{\pi(\epsilon_0^2 - 16f^2)N}} \cdot (1 \pm \cos(2fN(\phi - \delta))) \,, \tag{99}$$

$$\left(\frac{\partial P(\{n_L, n_R, (\mathbf{z}_L, \mathbf{z}_R)\})}{\partial \phi}\right)^2 \approx \left(\frac{P(n_L, n_R)}{2^{2(N-n_L-n_R)}}\right)^2 \cdot 8f^2N^2 \cdot \left(\frac{16\epsilon_0}{\pi(\epsilon_0^2 - 16f^2)N}\right) \cdot \sin^2(2fN(\phi - \delta)). \quad (100)$$

It follows from Equation (68) that

$$F(\{n_L, n_R, (\mathbf{z}_L, \mathbf{z}_R)\}) \approx \frac{P(n_L, n_R)}{2^{2(N-n_L-n_R)}} \cdot 8f^2N^2 \cdot \sqrt{\frac{16\epsilon_0}{\pi(\epsilon_0^2 - 16f^2)N}} \cdot \left(\frac{\sin^2(2fN(\phi - \delta))}{1 \pm \cos(2fN(\phi - \delta))}\right) \quad (101)$$

$$= \frac{P(n_L, n_R)}{2^{2(N-n_L-n_R)}} \cdot 8f^2N^{3/2} \cdot \sqrt{\frac{16\epsilon_0}{\pi(\epsilon_0^2 - 16f^2)}} \cdot (1 \mp \cos(2fN(\phi - \delta))). \quad (102)$$

Summing up these Fisher information terms over $\mathcal{Z}^{(n_L, n_R)}$ and using the bound from Equation (84) gives us

$$\sum_{(\mathbf{z}_L, \mathbf{z}_R) \in \mathcal{Z}^{(n_L, n_R)}} F(\{n_L, n_R, (\mathbf{z}_L, \mathbf{z}_R)\}) \gtrapprox P(n_L, n_R) \cdot 8f^2 \cdot \sqrt{\frac{16\epsilon_0}{\pi(\epsilon_0^2 - 16f^2)}} \cdot N^{3/2} \times \left(1 - 2\exp\left(-\frac{(\beta - 2\gamma)^2}{\epsilon_0 + 2\beta}N\right)\right) \times (1 - |\cos(2fN(\phi - \delta))|). \quad (103)$$

It follows that the full Fisher information will be bounded by

$$F = \sum_{n_L, n_R} \sum_{(\mathbf{z}_L, \mathbf{z}_R) \in \mathcal{Z}^{(n_L, n_R)}} F(\{n_L, n_R, (\mathbf{z}_L, \mathbf{z}_R)\}) \gtrapprox 8f^2 \cdot \sqrt{\frac{16\epsilon_0}{\pi(\epsilon_0^2 - 16f^2)}} \cdot N^{3/2} \cdot (1 - |\cos(2fN(\phi - \delta))|) \sum_{n_L, n_R} P(n_L, n_R). \quad (104)$$

D. Bounding the Probability of Typical (n_L, n_R)

It remains only to approximate $\sum_{n_L, n_R} P(n_L, n_R)$. Again, we lower bound this sum by considering only (n_L, n_R) pairs in the range defined by Equations (70) and (71). Define the interval

$$\mathcal{I}_{\beta,N} := \left[\frac{1}{2}\left(1 - \frac{\epsilon_0}{2}\right)N - \beta N,\ \frac{1}{2}\left(1 - \frac{\epsilon_0}{2}\right)N + \beta N\right].$$

Because n_L and n_R correspond to random time instances where the EPS photon definitively arrives at the left or right antenna, respectively, we see that

$$P(n_L, n_R) = \frac{1}{2}P_{B\left(\frac{N}{2}-d,\ 1-\frac{\epsilon_0}{2}\right)}(n_L) \cdot P_{B\left(\frac{N}{2}+d,\ 1-\frac{\epsilon_0}{2}\right)}(n_R) + \frac{1}{2}P_{B\left(\frac{N}{2}+d,\ 1-\frac{\epsilon_0}{2}\right)}(n_L) \cdot P_{B\left(\frac{N}{2}-d,\ 1-\frac{\epsilon_0}{2}\right)}(n_R), \quad (105)$$

where $B(n, p)$ represents a binomial distribution with n trials and success probability p (in this case, the probability that no source photon arrives or that both the source

and EPS photon arrive at the same antenna). The EPS is a superposition of two configurations: The first term of the sum in Equation (105) represents the configuration where $\frac{N}{2} - d$ EPS photons go to the left antenna and $\frac{N}{2} + d$ go to the right, and the second term represents the opposite configuration.

If we let $F_X(k; n, p)$ denote the cumulative distribution function of a random variable $X \sim B(n, p)$, that is,

$$F_X(k; n, p) = P(X \leq k),\ X \sim B(n, p), \tag{106}$$

then we have

$$\begin{aligned}
&\sum_{n_L, n_R \in \mathcal{I}_{\beta,N}} P(n_L, n_R) \\
&\quad = \frac{1}{2}\left[\sum_{n_L \in \mathcal{I}_{\beta,N}} P_{B\left(\frac{N}{2}-d,\ 1-\frac{\epsilon_0}{2}\right)}(n_L)\right] \cdot \left[\sum_{n_L \in \mathcal{I}_{\beta,N}} P_{B\left(\frac{N}{2}+d,\ 1-\frac{\epsilon_0}{2}\right)}(n_R)\right] \\
&\qquad + \frac{1}{2}\left[\sum_{n_L \in \mathcal{I}_{\beta,N}} P_{B\left(\frac{N}{2}+d,\ 1-\frac{\epsilon_0}{2}\right)}(n_L)\right] \cdot \left[\sum_{n_L \in \mathcal{I}_{\beta,N}} P_{B\left(\frac{N}{2}-d,\ 1-\frac{\epsilon_0}{2}\right)}(n_R)\right] \qquad (107) \\
&\quad = \frac{1}{2}\left[1 - 2 \cdot F_X\left(\left(1-\frac{\epsilon_0}{2}\right)\cdot\frac{N}{2} - \beta N;\ \frac{N}{2} - d,\ 1-\frac{\epsilon_0}{2}\right)\right] \\
&\qquad\qquad \times \left[1 - 2 \cdot F_X\left(\left(1-\frac{\epsilon_0}{2}\right)\cdot\frac{N}{2} - \beta N;\ \frac{N}{2} + d,\ 1-\frac{\epsilon_0}{2}\right)\right] \\
&\qquad + \frac{1}{2}\left[1 - 2 \cdot F_X\left(\left(1-\frac{\epsilon_0}{2}\right)\cdot\frac{N}{2} - \beta N;\ \frac{N}{2} + d,\ 1-\frac{\epsilon_0}{2}\right)\right] \\
&\qquad\qquad \times \left[1 - 2 \cdot F_X\left(\left(1-\frac{\epsilon_0}{2}\right)\cdot\frac{N}{2} - \beta N;\ \frac{N}{2} - d,\ 1-\frac{\epsilon_0}{2}\right)\right] \qquad (108) \\
&\quad = \left[1 - 2 \cdot F_X\left(\left(1-\frac{\epsilon_0}{2}\right)\cdot\frac{N}{2} - \beta N;\ \frac{N}{2} + fN,\ 1-\frac{\epsilon_0}{2}\right)\right] \\
&\qquad\qquad \times \left[1 - 2 \cdot F_X\left(\left(1-\frac{\epsilon_0}{2}\right)\cdot\frac{N}{2} - \beta N;\ \frac{N}{2} - fN,\ 1-\frac{\epsilon_0}{2}\right)\right]. \qquad (109)
\end{aligned}$$

Hoeffding's inequality states that

$$F_X(k; n, p) \leq \exp\left(-2n\left(p - \frac{k}{n}\right)^2\right), \tag{110}$$

provided that $k < np$. We can use this to bound Equation (109) as

$$\begin{aligned}
\sum_{n_L, n_R \in \mathcal{I}_{\beta,N}} P(n_L, n_R) &\geq \left[1 - 2 \cdot \exp\left(-\left(1-\frac{\epsilon_0}{2}\right)N \cdot \left(\left(1-\frac{\epsilon_0}{2}\right) - \frac{\left(1-\frac{\epsilon_0}{2}\right)\cdot\frac{1}{2} - \beta}{\frac{1}{2}+f}\right)^2\right)\right] \\
&\quad \times \left[1 - 2 \cdot \exp\left(-\left(1-\frac{\epsilon_0}{2}\right)N \right.\right. \\
&\qquad \left.\left. \cdot \left(\left(1-\frac{\epsilon_0}{2}\right) - \frac{\left(1-\frac{\epsilon_0}{2}\right)\cdot\frac{1}{2} - \beta}{\frac{1}{2}-f}\right)^2\right)\right], \qquad (111)
\end{aligned}$$

which is satisfied if $\beta > f \cdot \left(1 - \frac{\epsilon_0}{2}\right)$. This bound tends to 1 as N grows large. We can update our asymptotic lower bound from Equation (104) as

$$F \gtrapprox 8f^2 \cdot \sqrt{\frac{16\epsilon_0}{\pi(\epsilon_0^2 - 16f^2)}} \cdot N^{3/2} \cdot (1 - |\cos(2fN(\phi - \delta))|), \tag{112}$$

and we see that the envelope of this bound for a fixed value of f is equal to

$$F_{\text{env}} := 8f^2 \cdot \sqrt{\frac{16\epsilon_0}{\pi(\epsilon_0^2 - 16f^2)}} \cdot N^{3/2}, \tag{113}$$

causing our Cramér-Rao bound on the variance of our estimate of ϕ to scale as $\frac{1}{N^{3/2}}$ and our subsequent lower bound on standard deviation to scale as $\frac{1}{N^{3/4}}$ thereby beating the shot noise scaling of $1/\sqrt{N}$. In Figure 6, we plot the envelope of the Fisher information from Equation (113) for $\epsilon_0 = 0.05$ and several values of f. We see that even when f is well below $\frac{\epsilon_0}{4}$, the maximum allowable value in our analysis F_{env} quickly transitions from scaling on par with shot noise ($\sim N$) to the more desirable scaling of $\sim N^{3/2}$.

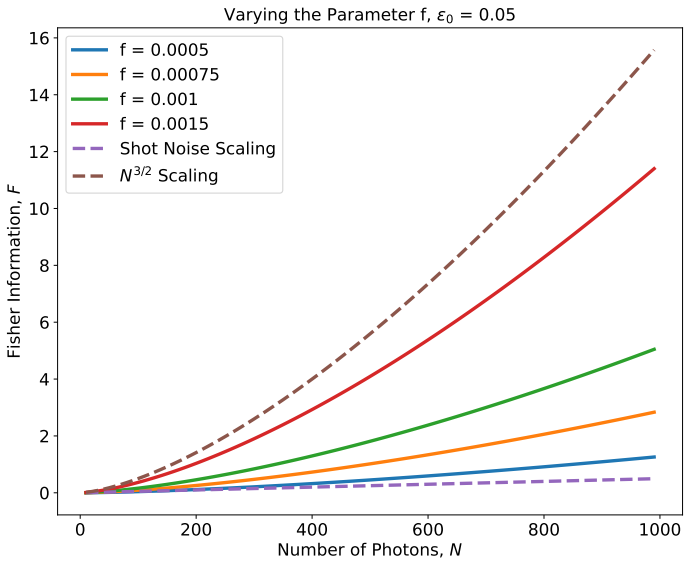


Figure 6. The scaling of the Fisher information in our quantum VLBI setup as the parameter f is varied but kept constant over the number of photons N. As f is raised toward its maximum value of $\epsilon_0/4 = 0.0125$, the Fisher information shifts away from the shot noise scaling of $\sim N$ toward a better scaling of $\sim N^{3/2}$.

Remark 3: From our previous observation (Remark 2) about the signs of $Z_1^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R)$ and $Z_2^{(n_L,n_R)}(\mathbf{z}_L, \mathbf{z}_R)$, summed terms $F(\{n_L, n_R, (\mathbf{z}_L, \mathbf{z}_R)\})$ will tend to either 1) all have a factor of $1 - \cos(2fN(\phi - \delta))$, 2) all have a factor of $1 + |\cos(2fN(\phi - \delta))|$, or 3) approximately half will have the first factor and half will have the second, causing the cosine terms to roughly cancel out in the summation. The result is that, in any of these cases, the envelope in Equation (113) will grow by a factor of 2.

In fact, if we select a constant k and allow the parameter f to vary as a function of N as

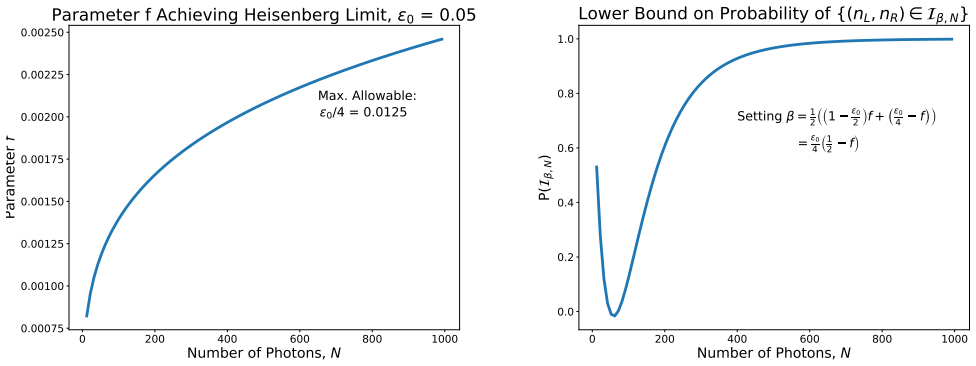
$$f \approx \sqrt{\frac{-\pi k N + \sqrt{\pi^2 k^2 N^2 + 64\pi\epsilon_0^3 k N}}{512\epsilon_0}}, \tag{114}$$

which is less than $\frac{\epsilon_0}{4}$ because the square root function is concave and increasing, then we can see that we can produce an envelope that is approximately

$$F_{env} \approx \frac{\sqrt{k}}{2} N^2, \tag{115}$$

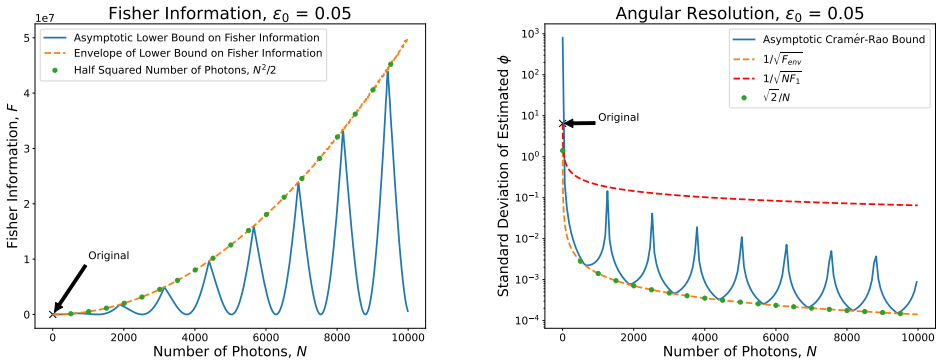
giving us a Cramér-Rao bound of $\frac{1}{F} \sim \frac{1}{N^2}$ and a lower bound on the standard deviation error scaling as $\sim \frac{1}{N}$, allowing us to achieve Heisenberg-limited scaling.

In Figure 7(a), we plot the value of f that achieves Heisenberg-limited scaling, setting $k = 10^{-9}$ in Equation (114) for $\epsilon_0 = 0.05$, and see that it stays well below the $\frac{\epsilon_0}{4}$ limit from Equation (93). In Figure 7(b), we plot the corresponding lower bound from Equation (111) on the probability of the typical set of events $\mathcal{I}_{\beta,N}$ considered in our analysis. We set β equal to the midpoint of $\left[\frac{\epsilon_0}{4} - f,\ \left(1 - \frac{\epsilon_0}{2}\right) \cdot f\right]$, its allowable range as determined by Equations (70), (71), and (111). As we can see, the likelihood of detection signatures falling in $\mathcal{I}_{\beta,N}$ quickly approaches 1 with increasing N. In Figure 8(a), we plot the resulting value of the lower bound on the Fisher information given by Equation (112), and we see how it lies along the envelope described by Equation (113), which exhibits Heisenberg-limited scaling. The corresponding angular resolution for estimating ϕ from Equation (4) is shown in Figure 8(b), showing that we quickly observe improvement in our estimation uncertainty with increasing N when compared to the original setting of $N = 1$—the single-photon EPS case from Gottesman et al. [3]. The figure also depicts the angular resolution improvement over N repeated applications of the single-photon EPSs, which lies at the shot noise limit.



(a) **(b)**

Figure 7. (a) The value of the parameter f needed to achieve Heisenberg-limited scaling of the Fisher information, setting $k = 10^{-9}$ in Equation (114). We see that f stays well below its maximum allowable limit of $\epsilon_0/4$. (b) The associated probability of the set typical set $\mathcal{I}_{\beta,N}$ considered in our analysis, setting β equal to the midpoint of $\left[\frac{\epsilon_0}{4} - f,\ \left(1 - \frac{\epsilon_0}{2}\right) \cdot f\right]$, its allowable range.



(a) **(b)**

Figure 8. (a) The asymptotic lower bound on Fisher information from Equation (112) obtained by choosing the parameter f as in Equation (114), setting $k = 10^{-9}$. We see that its envelope from Equation (113) approaches Heisenberg-limited scaling of $\sim N^2$. (b) The corresponding Cramér-Rao bounds on angular resolution resulting from the asymptotic bound on F and its envelope. The original single-photon EPS case from Gottesman et al. [3] is annotated in each figure, as well as the angular resolution achievable from repeating it N times ($1/\sqrt{NF_1}$, where F_1 is defined in Equation (29)).

VII. Concluding Remarks and Future Directions

While we have demonstrated that the point-source angular resolution of VLBI can be enhanced by sharing multiphoton quantum states between antennas, there are several major obstacles to implementing such a system in practice. The quantum states we have employed are exotic, and lie beyond the realm of what has been implemented to date. Indeed, even the NOON states we employed in Section III have only been realized experimentally for up to $N = 5$ [13]. Furthermore, the VLBI setup of Gottesman et al. [3] (depicted in Figure 2) requires quantum memory to interface more than two apertures, and our concept would have much more demanding quantum memory requirements. Such technology is relatively new, though in active development.

We remark that our arguments have assumed rather capable photon-number resolution (PNR) for our detectors, particularly in the case of all source and EPS photons arriving in a single time bin. In the case where our source is weak, with at most a single photon arriving within a time bin, this becomes less of an issue, though one may still ask how our results would be affected by non-ideal PNR measurements. This would require us to recompute our approximations for the Fisher information assuming modeled imperfect measurement outcomes with a modeled probability distribution, which we leave for future work. Likewise, we could also seek to compute the effect of extra received photons due to noise. In some sense, the EPSs we designed in Equation (31) resemble error-correcting codes, where we have considered errors to be events in which the location of an EPS photon is determined. By the same token, if we include error events where multiple non-EPS photons arrive in a single time bin (say from both a source and from noise), our system should be somewhat robust. More

interesting would be accounting for cases where only a noise photon arrives and may be potentially confused for a source photon. We again leave this analysis for the future.

References

[1] A. R. Thompson. Fundamentals of radio interferometry. *Synthesis Imaging in Radio Astronomy II, A Collection of Lectures from the Sixth NRAO/NMIMT Synthesis Imaging Summer School. Edited by G. B. Taylor, C. L. Carilli, and R. A. Perley. ASP Conference Series*, 180, 1999.

[2] B. Campbell. So you want to do vlbi. *Fourth EVN VLBI School*, November 1999.

[3] D. Gottesman, T. Jennewein, and S. Croke. Longer-baseline telescopes using quantum repeaters. *Physical Review Letters*, 109:070503, August 2012.

[4] V. Giovannetti, S. Lloyd, and L. Maccone. Quantum-enhanced measurements: Beating the standard quantum limit. *Science*, 306:1330–1336, November 2004.

[5] P. Kok, S.-L. Braunstein, and J.-P. Dowling. Quantum lithography, entanglement and heisenberg-limited parameter estimation. *Journal of Optics B: Quantum and Semiclassical Optics*, 6(8):S811, July 2004.

[6] M. J. Holland and K. Burnett. Interferometric detection of optical phase shifts at the heisenberg limit. *Physical Review Letters*, 71(9):1355, August 1993.

[7] R. A. Campos, C. C. Gerry, and A. Benmoussa. Optical interferometry at the heisenberg limit with twin fock states and parity measurements. *Physical Review A*, 68(2):023810, August 2003.

[8] M. D. Lang and C. M. Caves. Optimal quantum-enhanced interferometry. *Physical Review A*, 90:025802, 2014.

[9] M. D. Lang and C. M. Caves. Optimal quantum-enhanced interferometry using a laser power source. *Physical Review Letters*, 111(17):173601, October 2013.

[10] G. S. Thekkadath, M. E. Mycroft, B. A. Bell, C. G. Wade, A. Eckstein, D. S. Phillips, R. B. Patel, A. Buraczewski, A. E. Lita, T. Gerrits, S. W. Nam, M. Stobińska, A. I. Lvovsky, and I. A. Walmsley. Quantum-enhanced interferometry with large heralded photon-number states. *npj Quantum Information*, 6(89), October 2020.

[11] Mankei Tsang. Quantum nonlocality in weak-thermal-light interferometry. *Physical Review Letters*, 107(27):270402, December 2011.

[12] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.

[13] I. Afek, O. Ambar, and Y. Silberberg. High-NOON states by mixing quantum and classical light. *Science*, 328(5980):879–881, 2010.